Appendix A

The Tsai Camera Model

The Tsai camera model [62] describes a camera as a pinhole projector combined with radial lens distortion and is completely defined by 12 parameters:

- (3) 3D rotation
- (3) 3D translation
- (1) focal length
- (2) lens distortion
- (1) aspect ratio
- (2) image center

Tsai observed that lens distortion is usually modeled well with only one parameter, and so the actual model used has 11 parameters.

To project a 3D world point $\bar{p}_w$ into an image, the 3D coordinate is first rotated and translated into camera coordinates, yielding $\bar{p}_c$:

$$\bar{p}_c = R\bar{p}_w + \bar{T} \quad \text{where} \quad R = R_{\theta_z}R_{\theta_y}R_{\theta_x} \quad \bar{T} = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad \bar{p}_\alpha = \begin{bmatrix} X_\alpha \\ Y_\alpha \\ Z_\alpha \end{bmatrix} \tag{A.1}$$

R_{θ_k} is a 3x3 rotation matrix, rotating about coordinate axis k by angle θ_k, and T_k is a translation along coordinate axis k. The six camera parameters used here, θ_k and T_k, are collectively referred to as extrinsic parameters. After this 3D transformation, $\vec{p}_c$ is perspectively projected into undistorted sensor coordinates (x_u, y_u), using the focal length f:

$$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = \frac{f}{Z_c}\begin{bmatrix} X_c \\ Y_c \end{bmatrix} \tag{A.2}$$

Next, the sensor coordinates are radially distorted, using the distortion parameter κ_1, to acquire distorted sensor coordinates (x_d, y_d) :

$$\begin{bmatrix} x_u \\ y_u \end{bmatrix} = (1 + \kappa_1 r^2)\begin{bmatrix} x_d \\ y_d \end{bmatrix} \qquad r^2 = x_d^2 + y_d^2 \tag{A.3}$$

Note that this equation is formulated as an inverse mapping; solving for distorted coordinates requires the solution of a cubic polynomial. The image coordinates (x_f, y_f) are computed by applying the aspect ratio a and image center (C_x, C_y) :

$$\vec{q} = \begin{bmatrix} x_f \\ y_f \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_d \\ y_d \end{bmatrix} + \begin{bmatrix} C_x \\ C_y \end{bmatrix} \tag{A.4}$$

This model degenerates into the general camera model of Equation (2.4) when there is no lens distortion (i.e., $\kappa_1 = 0$), and further degenerates into the simple projection of Equation (2.1) where there also is no rotation ($R = I$), no translation along the Z axis ($T_z = 0$), equal focal lengths across all cameras, a unity aspect ratio ($a = 1$), and an image center (C_x, C_y) at (0,0).